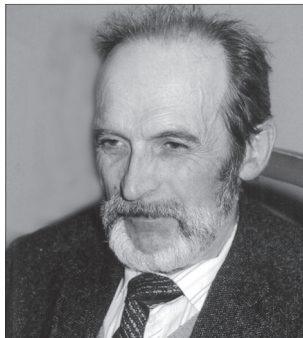




АСТАХОВА
Татьяна Сергеевна
Старший научный сотрудник
ВИНИТИ РАН
✉ typo@viniti.ru



БЕЛООЗЕРОВ
Виктор Николаевич
Кандидат филологических
наук, доцент, заведующий
сектором ВИНИТИ РАН
✉ systemling@narod.ru



СУРИКОВА
Надежда Германовна
Научный сотрудник
ВИНИТИ РАН
✉ typo@viniti.ru

ДМИТРИЕВА
Елена Юрьевна
Кандидат технических наук,
заведующий отделением
ВИНИТИ РАН
✉ typo@viniti.ru

УДК 025.45.05УДК-028.27

ЭЛЕКТРОННОЕ ИЗДАНИЕ УНИВЕРСАЛЬНОЙ ДЕСЯТИЧНОЙ КЛАССИФИКАЦИИ

Описывается структура и использование нового информационного продукта, содержащего полную эталонную версию таблиц УДК на русском языке для широкого распространения. Издание будет полезно систематизаторам научных публикаций в библиотеках, информационных центрах и издательствах.

УДК: книжные и электронные издания

Универсальная десятичная классификация (УДК) является наиболее развитой международно принятой классификационной системой для тематического описания документов. В нашей стране она утверждена как обязательная для индексирования тематики всей издаваемой литературы (наряду с

ББК) и всех документов в автоматизированных системах научно-технической информации. Ответственность за содержание классификационных таблиц УДК и право их распространения принадлежит международному Консорциуму УДК (UDC Consortium, The Hague, Netherlands)¹. Россия представлена в Управляющем совете консорциума директором ВИНИТИ академиком Ю. М. Арским, а ВИНИТИ обладает исключительным правом и обязанностью издавать и распространять материалы УДК на русском языке. В соответствии с этим правом и обязанностью, ВИНИТИ в 2001 – 2009 годах осуществил полное издание УДК в 10 томах (12 книгах) [1] и теперь периодически издаёт бюллетень

¹ Официальный сайт Консорциума УДК: <http://www.udcc.org>.

изменений и дополнений к таблицам². Изданы также методические руководства и практические пособия по использованию УДК [3] [4].

Изданные тома распространяются как в бумажном виде, так и в электронном – в виде оригинал-макетов в формате Word. Однако необходимость встраивать таблицы УДК в пользовательские автоматизированные системы породила потребность в издании таблиц в простой текстовой форме, где все элементы классификационных записей размечены легко опознаваемыми однозначными пометами. Соответствующий массив классификационных записей подготовлен в 2013 г. для издания на компакт-дисках. Этот массив содержит актуализированный эталон российского издания УДК. Он сформирован на основе десяти томов 4-го полного издания таблиц УДК на русском языке, в которые были в иерархической последовательности внесены все обновления, опубликованные ранее в отдельных выпусках изменений и дополнений. Кроме этого, было проведено дополнительное редактирование всего информационного массива для устранения некорректного представления данных и более глубокого системного структурирования. Таким образом, на рынке информационных продуктов появляется объединенный **Эталонный массив таблиц УДК**, включающий все изменения, дополнения и редакционные правки. Предполагается в дальнейшем поддерживать этот массив в актуальном состоянии, вводя в него обновления вслед за принятием их в Консорциуме УДК и получением соответствующего выпуска бюллетеня Extensions and Corrections to the UDC. Таким образом, мы будем иметь так называемое **обновляемое электронное издание УДК**, и каждый приобретатель его получит таблицы УДК, которые действительно на момент приобретения. По желанию потребителей на основе полного эталонного массива могут быть сформированы тематические фрагменты.

² Последний выпуск бюллетеня см. [2].

Структура продукта

Эталонный массив таблиц УДК содержит перечень полей, упорядоченный в соответствии с иерархической структурой классификационных таблиц. Он включает основные тематические классы по всем отраслям знания, а также классы вспомогательных таблиц (общие и специальные определители), которые выполняют функции стандартных классификационных фасет.

Физически массив представляет собой текстовый файл, в котором каждое поле классификационных записей занимает отдельную строку, начинающуюся с функциональной метки поля из числа указанных в таблице 1.

Массив текстовых строк делится на классификационные записи четырёх типов: основной класс, исключённый

Ключевые слова:

УДК, эталонные таблицы, библиографическая классификация, информационный продукт, электронное издание.

Keywords:

UDC, reference tables, bibliographic classification, information product, electronic edition

Таблица 1.
Функциональные метки полей

Метка поля	Наименование поля
<01>	иерархический код класса (индекс УДК)
<02>	наименование класса
<03>	подзаголовок класса (резервное поле)
<04>	аспектная ссылка (отсылка)
<05>	методические указания
<06>	примеры комбинированных индексов
<07>	указание на подразделения по аналогии (подразделять как...)
<08>	примеры, относящиеся к полю <07>
<09>	ссылки на другие классы
<01И>	иерархический код (индекс) исключенного класса
<опро1>	иерархический код (индекс) определителя
<опро2>	наименование класса определителя
<опро3>	подзаголовок определителя (резервное поле)
<опро4>	аспектная ссылка (отсылка)
<опро5>	методические указания
<опро6>	примеры комбинированных индексов
<опро7>	указание на подразделения по аналогии (подразделять как...)
<опро8>	примеры, относящиеся к полю <опро7>
<опро9>	ссылки на другие классы определителей
<опро1И>	иерархический код (индекс) исключенного определителя

преимущественным отличием русской версии УДК является ее большая полнота

класс, специальный определитель, исключённый специальный определитель. Запись основного класса начинается со строки поля <01>, за которой должно следовать поле <02> и могут далее идти поля <03> – <09>. Поля <03> и <07> могут быть представлены только одним экземпляром, а остальные поля могут повторяться. Запись специального определителя начинается строкой поля <опр01>, за которой в том же порядке следуют поля, аналогичные записи основного класса, указанные в таблице 1. Записи исключённого класса и исключённого определителя начинаются строкой поля <01И> и <опр01И> соответственно, за которой должна следовать строка <02>, и может следовать строка <09>. Записи общих определителей не снабжены специальными метками; они опознаются только по особенностям иерархического кода, который помещается в строке с меткой <02>. Все строки, заключённые между двумя смежными «начальными» строками (<01>, <опр01>, <01И>, <опр01И>), относятся к записи класса, код которого содержится в предыдущей «начальной» строке.

За основу формирования эталонного массива была взята структура англоязычной стандартной (средней) версии УДК «Master reference file», предлагаемой своим пользователям Международным Консорциумом УДК³.

Преимущественным отличием русской версии является ее большая полнота. Российский эталон соответствует так называемой «полной» версии УДК, классы которой в ряде случаев развиты вглубь больше, чем в «стандартной»

версии. Кроме того здесь представлена информация об исключённых классах, которой нет в «Master reference file», а ссылки на другие классы представлены более подробно (включая наименования классов). Так что объем нашего продукта превышает англоязычный аналог почти в два раза.

Поисковые возможности

Просмотр и поиск элементов Эталонного массива таблиц УДК удобно реализовать средствами текстового редактора Microsoft Word (желательно версии Word 2007 или более поздней). Он позволяет осуществлять поиск по всем возможным словоформам (по ключевым словам, с усечением, по отдельным фразам) в начале, в середине строки, а также по кодам классов (индексам).

Для поиска в редакторе Word открывается диалоговое окно Найти и заменить. В командную строку диалогового окна, в зависимости от задачи, стоящей перед систематизатором, заносятся необходимые символы: а) цифровые – при поиске по коду класса, б) буквенные – при поиске по содержанию класса.

Рассмотрим варианты поисковых процедур.

А. Для поиска записи определённого класса в командную строку следует ввести точное значение кода класса или точный текст наименования класса с предварительной меткой соответствующего поля.

При вводе кода класса программа выводит на экран начальную строку искомого класса, содержащую заданный код. Ниже этой строки располагается все имеющиеся поля данного класса. Ещё ниже расположены все подклассы искомого класса. Нажатие мышки позволяет переходить от одного подкласса к другому и так просмотреть все релевантные данные.

Например, по запросу <01>612.8, где <01> – метка поля иерархического кода класса, находим наименование класса **Нервная система. Органы чувств.**

³ Подробнее о продукте см. на <http://www.udcc.org/index.php/site/page?view=mrf>

Другие поля в этом классе не предусмотрены, но ниже на экране следуют классы специальных определителей, применяемые здесь, и большой ряд подклассов, соответствующих разделам нервной системы.

В другом случае по запросу **<01> 615.479.5** выдаётся наименование класса **Одежда для пациентов** и ещё два поля классификационной записи: ссылка на другой класс **<09>615.477 Ортопедический материал. Резиновые изделия и т. п.** указание на подразделение по аналогии **<07> 615.479.5.подразделять как 615.479.4**. Подклассы у данного класса отсутствуют в классификационной таблице, но они могут быть образованы согласно этому указанию на этапе применения УДК.

Б. При вводе точного текста наименования класса код класса располагается на строке выше выделенного поисковой программой текста. Остальные поля записи – ниже. При этом может быть найдено кроме искомого класса и другие, наименования которых в начале совпадают с заданным. Так при запросе **<01> Головные уборы** кроме класса **646.5 Головные уборы** из раздела домашнего хозяйства выдаются также классы:

- **615.479.41 – Головные уборы, шапочки** – из медицины
- **687.4 – Головные уборы. Производство головных уборов** – из лёгкой промышленности
- **687.463 – Головные уборы для мальчиков** –
- **687.465 – Головные уборы для девочек** – и т. д.

В данном случае эта информация является нерелевантной.

В. В случае потребности найти подходящий класс для заданной темы «Головные уборы» данные, указанные выше, были бы полезны. Но тогда следует запрос оформлять как *запрос по ключевой слову*, т. е. без метки поля. Тогда программа выдаст все классы, где заданное выражение встречается в любом поле и в любом месте текста, а не только с начала строки. А именно, будут дополнительно выданы классы:

- **687.43 – Мужские головные уборы в целом. Мужские шапки** –
- **687.433 – Специальные головные уборы для защиты от солнца. Тропические шлемы. Пробковые шлемы** –
- **929.628.2/.5 – Шлемы и головные уборы** – и др.

Другой пример. По запросу электронный беспорядок в тексте находится и выделяется эта фраза в единственной строке:

<02>Электронный беспорядок (нерегулярности в структуре электронных оболочек атомов или ионов). Электронные дефекты, которая по вышестоящей строке **<опр01>544.022.37** опознаётся как специальный определитель в классе физической химии.

В данном случае получается, что на запрос по ключевому слову выдана одна классификационная запись, т.е. информационный шум сведен к нулю. Однако в общем случае поиск может выдать записи из разных разделов классификации, часть которых могут оказаться не pertinentными.

Г. При потребности найти все классификационные записи, имеющие ссылку на данный класс, следует в командной строке набирать код класса с меткой **<09>**. Например, на запрос **<09>616.893** будет выдана строка **<09>616.893 Острое помешательство**, которая является ссылкой в записи специального определителя **616.89-02 Этиология психозов**.

Д. Примеры комбинированных рубрик приводятся в строках с меткой **<06>**, которые входят в состав записи одного из классов, участвующих в данной комбинации. Эту метку следует включать в запрос для поиска тех комбинированных индексов, в которых код заданного класса стоит на первом месте. Например, запрос **<06>745.9** приведёт к выдаче строк **<06>745.9.05:27**, **<06>745.9.05:394.4**, **<06>745.9.05:792.023**, которые являются примерами комбинаций класса **745.9 Цветочное искусство** со специальным



определителем применения .05 и классами из других разделов классификации.

Е. Если же нужно найти примеры, в индексах которых заданный класс стоит не на первом месте, в запросе нужно набирать не метку поля, а символ присоединения кода заданного класса к индексу. Например, поиск по запросу : **621.326** приводит, в частности, к выдаче строки **<06>778.23:621.326 Лампы накаливания для проекторов**, которая входит как пример в запись класса **778.23 Источники света для проекторов**.

Так же следует составлять запрос при поиске комбинаций с общими и специальными определителями, т. е. определители нужно вводить с их специальными символами и без метки поля. Так на запрос -053.6 будут выданы примеры индексов, относящихся к воздействию разных факторов на подростков и, в частности, строка **<06>791.43-053.6**, которая входит в состав записи класса **791.43 Воздействие кино на общество**.

Пример составления сложного индекса для названия статьи

Поиск в эталонном массиве УДК позволяет составлять подробные многоаспектные индексы, описывающие публикацию с разных сторон. Рассмотрим в качестве примера статью в журнале со следующим заглавием:

Тактико-технические характеристики океанского крейсера русского флота 80-х годов XIX века.

Ключевые слова:

- а) крейсер
- б) тактико-технические характеристики
- в) русский, российский
- г) XIX век
- д) 80-е годы.

При запросе «**крейсер**» выделяется несколько классов, учитывающих специфику данного термина. Здесь систематизатору нужно определиться, в каком аспекте употребляется это понятие. Выбираем класс **623.822.3 Крейсера** из раздела военной техники.

При запросе «**тактико-технические характеристики**» нас интересуют именно тактико-технические характеристики крейсеров. Получаем ответ: **623.822.3-9**, где определитель -9 найден в записи вышестоящего класса **623. Военная техника**.

При поиске по ключевому слову «**русский**» систематизатор должен употребить термин «**российский**», поскольку ясно, что в статье идёт речь не об этнической, а о государственной принадлежности корабля. Также систематизатор должен применить усечённую форму термина, чтобы учесть все формы словоизменения, которые могут присутствовать в тексте записей. Получаем определитель страны (**470**) **Россия**.

Для уменьшения информационного шума при запросе «**XIX**» добавляем элемент метки **>**. Получаем ответ: «**18**». Для обозначения десятилетий добавляем цифру 8 (по правилам составления индексов, получаем: «**188**»). Далее составляется сложный индекс по принятым правилам УДК и получаем требуемый сложный индекс **623.822.3-9(470)«188»**, адекватно описывающий содержание статьи.

Применение продукта

Предлагаемый продукт поможет каждому индивидуальному или коллективному пользователю



УДК, работающему с любым изданием или электронной версией УДК, сверить и откорректировать применяемые им классы (индексы) УДК, актуализировать свой информационно-поисковый аппарат. Это будет способствовать формированию единого информационного пространства в сегменте применения УДК как информационно-поискового языка. Кроме этого, жестко упорядоченная структура информационного массива может помочь начинающему пользователю освоить УДК как классификационную систему. Глубокий уровень структурной разметки поможет разным пользователям самостоятельно формировать БД УДК для решения своих профессиональных задач.

Издание будет полезно систематизаторам научных публикаций в библиотеках, информационных центрах и издательствах.

Источники

1. УДК. Универсальная десятичная классификация: Полное издание на русском языке. Тома 1 – 10 / Глав. ред. Ю. М. Арский, общ. ред. Т. С. Астахова. – М.: ВИНТИ РАН, 2001 – 2009.
2. УДК. Универсальная десятичная классификация: Изменения и дополнения. Выпуск 6 / Глав. ред. Ю. М. Арский, общ. ред. Т. С. Астахова. – М.: ВИНТИ РАН, 2013.

3. Учебное пособие по Универсальной десятичной классификации. Изд. 2-е, испр. и доп. / Науч. ред. В. Н. Белоозеров – М.: ВИНТИ РАН, 2009. – 188 с.

4. УДК. Универсальная десятичная классификация: Консультации по индексированию (Ответы на вопросы пользователей). Вып. 4 / В. Н. Белоозеров, О. В. Смирнова – М.: ВИНТИ РАН, 2012. – 136 с.

Tatiana S. ASTAKHOVA,

Senior Research Assistant, VINITI RAN

Viktor N. BELOOZEROV,

Candidate of Philological Sci., Associate Professor, Head of Division of VINITI RAN

Elena Yu. DMITRIEVA,

Candidate of Technical Sci., Head of VINITI RAN

Nadezhda G. SURIKOVA,

Research Assistant of VINITI RAN

Electronic Edition of UDC

The article provides the description of the structure and use of new information product containing the complete master version of UDC tables in Russian language for wide distribution. The edition will be useful to the systematicians of scientific papers in libraries, information centers and publishing houses.